VizThis : Rule-based Semantically Assisted Information Visualization

Owen Gilson[1], Nuno Silva[2], Phil W. Grant[1], Min Chen[1], João Rocha[2]
[1] Department of Computer Science, Swansea Univeristy, UK
[2] ISEP, Instituto Politécnico do Porto, Portugal

"Visualization as Mapping"

Information Visualization can be considered as the task of mapping data entities in the source file to representation artefacts in the target visualization format. As such, mapping plays a significant role in this process. The field of ***Ontology Mapping*** presents a wealth of work in the information mapping domain which we propose to exploit by applying it to Information Visualization of semantically rich information. The realisation of this approach is demonstrated in the software tool, VizThis.



Features

Data Analysis: The source is analysed to try to predicate the nature and structure of the data. [see Step 1]
Semantic Assistance: The AutoMap button uses the Analysis and Categorisation of the Source and Target to automatically predicate semantically valid mappings. [see Step 1]
Data Cleansing: The source data can have simple data cleansing operations applied (regular expressions or search and replace operations). [see Step 2]
Tweaking of Mappings: The mappings created by Semantic Assistance may not be the best mapping possible, so the user is able to tweak any mappings. [see Step 3]
Rules Warnings: The VizThis tool will alert users when the mappings are in an invalid state. For example, it is not possible to map multiple source entities to a single target artefact (unless concatenation is used). [see screenshot above]

Benefits

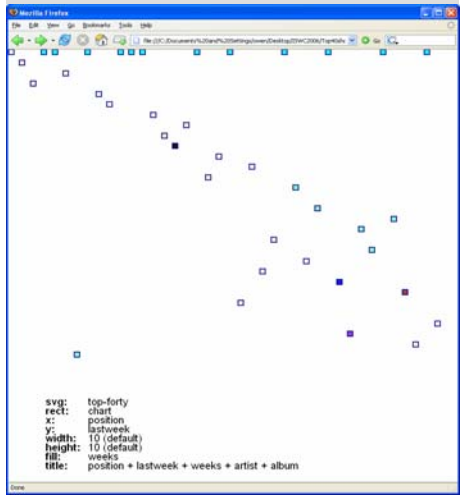
Automaticity: Much of the mapping process can be automated, or at least "semantically intelligent" guesses made.
Constraints: We can define the nature and behaviour of Source Entities and Target Artefacts in a way which allows us to derive rules. This allows us to constrain which entities can be mapped to which artefacts, thus simplifying the system and reducing users' work.
Generality: The technique is valid for any semantically rich mark-up language, whether it is tree or graph based, a number of different source and target formats from many domains can be supported.
Multimedia Target: Information can be represented in any format and therefore any media. So, the target is not limited to graphical representations, but also textual and audial too.

BBC Top 40 Music Chart

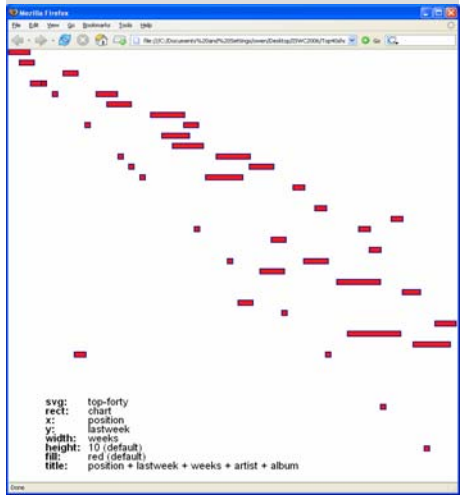
Real-life XML data source. Screen scraped from the BBC website.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<top-forty xmlns="http://www.w3.org/topforty">
    <chart position="1">
        <lastweek>1</lastweek>
        <weeks>(5)</weeks>
        <image>http://www.bbc.co.uk/radio1/media/images/
        artists/gnarlsbarkley/060323_cd_crazy_70.jpg</image>
        <artist>Gnarls Barkley</artist>
        <album>Crazy</album>
        <uri>http://www.gnarlsbarkley.com/</uri>
    </chart>
    …
</top-forty>
```

Step 1. Fully automatic – no human intervention



Step 2. Source data cleansed



Step 3. Source data cleansed / Mappings tweaked

